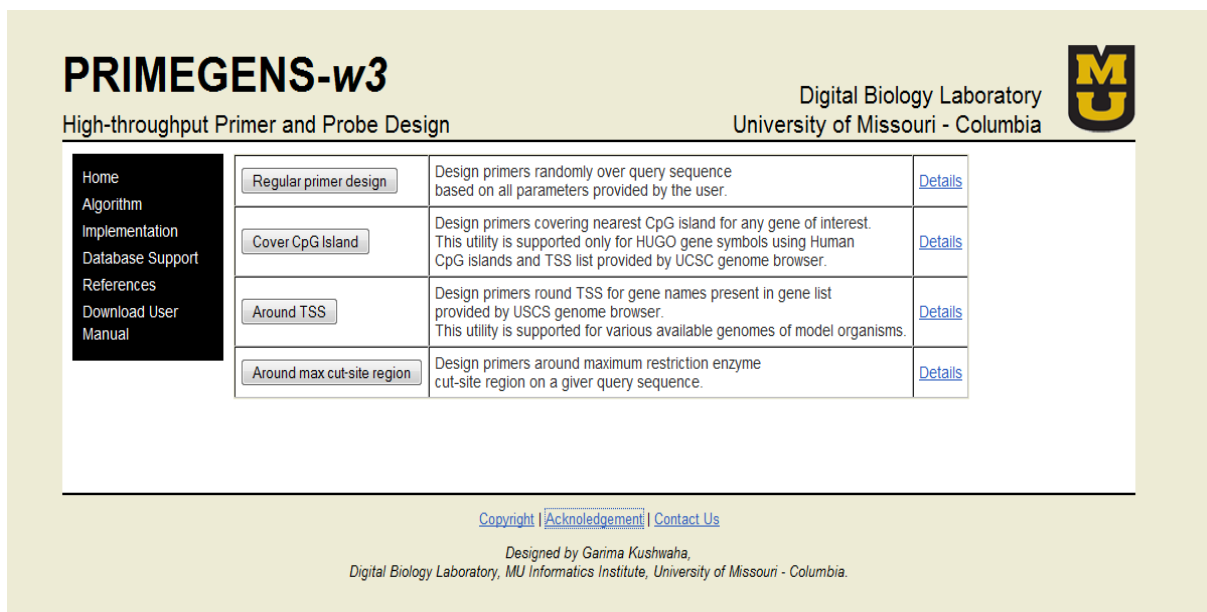


PRIMEGENSw3 User Manual

PRIMEGENSw3 is Web Server version of PRIMEGENS program to automate high-throughput primer and probe design. It provides three separate utilities to select targeted regions of interests from genome for PCR amplification long with its regular primer design process. PRIMEGENSw3's different utilities for primer and probe design are:

1. **Regular Primer Design.**
2. **Cover CpG Island.**
3. **Around TSS.**
4. **Around max cut-sit region.**

Figure 1 shows the webpage showing different options for the user choos for primer or probe design for these utilities.



PRIMEGENS-w3
High-throughput Primer and Probe Design

Digital Biology Laboratory
University of Missouri - Columbia

Home
Algorithm
Implementation
Database Support
References
Download User Manual

Regular primer design	Design primers randomly over query sequence based on all parameters provided by the user.	Details
Cover CpG Island	Design primers covering nearest CpG island for any gene of interest. This utility is supported only for HUGO gene symbols using Human CpG islands and TSS list provided by UCSC genome browser.	Details
Around TSS	Design primers round TSS for gene names present in gene list provided by USCS genome browser. This utility is supported for various available genomes of model organisms.	Details
Around max cut-site region	Design primers around maximum restriction enzyme cut-site region on a giver query sequence.	Details

[Copyright](#) | [Acknowledgement](#) | [Contact Us](#)

Designed by Garima Kushwaha,
Digital Biology Laboratory, MU Informatics Institute, University of Missouri - Columbia.

Figure 1: PRIMEGENSw3 page for choosing between different utilities to design primers or probes.

Cover CpG Island. “Primer design covering CpG Island” is one of the unique features of PRIMEGENS-v2, which can be used to study methylation patterns of various oncogenes and tumor suppressor genes. This feature designs primers for genes that have CpG islands present in close proximity to their respective TSSs. Primers can be designed to amplify genes whose expressions are suspected to be influenced by nearby CpG islands. Detailed description for working of this utility is present on website as “Details” link in front of the link to this utility.

Around TSS. “Primer design covering TSS” is a feature of PRIMEGENS, which helps the designed primers to cover the region around a transcription start site (TSS) of any gene. To cover specific region around the TSS of any gene, the user is only required to provide gene symbols for which the primer design is required. PRIMEGENS is capable of extracting their respective TSSs from the UCSC Genome Database (currently for March 2006 assembly). Detailed description for working of this utility is present on website as “Details” link in front of the link to this utility.

Around max cut-sit region. PRIMEGENS can also be used to search for regions with the maximum enzyme digestion sites (cut-sites) within each query sequence and design primers around these cut-sites. This ensures the presence of cut-sites in the PCR product and is very useful in Methylation-specific PCR. Detailed description for working of this utility is present on website as “Details” link in front of the link to this utility.

For each of these utilities, PRIMEGENSw3 has a simple sequence of operations, which consist of two basic steps: **1) Uploading data files** (PCR templates file for primer design and optional database for cross-hybridization check); **2) Primer design specifications** which consist of setting various design parameters (for example, Primer3 parameters, BLAST parameters for cross hybridization check, etc.); and **3) Program execution and result visualizations**. It allows user to select three different algorithms for primer design in each of its utility. They are 1) Sequence-specific Primer Design (SSPD), allowing primer design for any random DNA sequence; 2) Fragment-specific Primer Design (FSPD), allowing multiple primer pair design distributed uniformly across target sequence for investigating large sequences; 3) Probe-specific Primer Design (PSPD), allowing users to design target sequence-specific probes and associated primers pairs. In addition to this, it can also be used to design sequence-specific probes.

Using web server version of PRIMEGENS software is a three step process as follows:

Step 1: Upload Input files.

For Regular Primer Design.

To design primers and probes, PRIMEGENS require two types of inputs. One is the query file having the sequence for which primers/probes need to be designed and the database file having all the other sequence that are present in the PCR reaction. Sequences in database file are the sequences to which PRIMEGENS will check for any potential cross hybridization and thereby select primer/probe that are specific to the sequence of interest from sequence mixture.

On PRIMEGENSw3 web-server, user can upload the query sequence (PCR template) file and their own custom database file (sequence mixture in PCR) or use available genomes supported by PRIMEGENS. PRIMEGENSw3 do also provide different sample data for both query and database sequences for users to test primer/probe design using PRIMEGENS algorithms. As per their selection, the corresponding upload or selection box gets activated for the user to provide respective option.

If any of these files, query or database file is not uploaded by the user before hitting submit button, the program will exit giving the error message as “Query file has not been uploaded.” or “Database file has not been uploaded.”.

Figure 2(a-c) shows the webserver page having various options for input files required by PRIMEGENSw3. Figure2(c) shows the available genomes options on webserver.

(a)

PRIMEGENSw3
High-throughput Primer and Probe Design

Digital Biology Laboratory
University of Missouri - Columbia

Home
Algorithm
• SSPD
• PSPD
• FSPD
• Check Binding Specificity
• Probe Design Only
Implementation
Database Support
References
Download User Manual

Input Files

Upload Query file

☐ Use Sample File [View Sample files](#)

☐ Upload your own file [Browse...](#)

• If "no sequence" query format used, database must contain corresponding sequences.
• Use ">chrX:start-end" in "no sequence" query format for genomic database.
• Please check format from sample files [Details](#)

Upload Database file

☐ Use Sample File [View](#)

☐ Upload your own file [Browse...](#)

☒ Available Genomes

• Use S. Cerevisiae from available genomes, if Format-3/Format-4 query sample is used.

[Next](#) [Reset](#)

Note: This website has been tested on Mozilla Firefox and Internet Explorer 8.

Select “Available Genomes” option to select the option to upload the database for browsing your file.

(b)

PRIMEGENSw3
High-throughput Primer and Probe Design

Digital Biology Laboratory
University of Missouri - Columbia

Home
Algorithm
• SSPD
• PSPD
• FSPD
• Check Binding Specificity
• Probe Design Only
Implementation
Database Support
References
Download User Manual

Input Files

Upload Query file

☒ Use Sample File [View Sample files](#)

☐ Upload your own file [Browse...](#)

• If "no sequence" query format used, database must contain corresponding sequences.
• Use ">chrX:start-end" in "no sequence" query format for genomic database.
• Please check format from sample files [Details](#)

Upload Database file

☐ Use Sample File [View](#)

☐ Upload your own file [Browse...](#)

☒ Available Genomes

• Use S. Cerevisiae from available genomes, if Format-3/Format-4 query sample is used.

[Next](#) [Reset](#)

Note: This website has been tested on Mozilla Firefox and Internet Explorer 8.

After uploading two corresponding files, click next to move on to next page

Figure 2: Input file page for Regular Design Utility.

For Around CpG Design.

Figure 3 shows the input file page for Around TSS utility. Here, the query file is Gene symbol list file. The gene symbols are taken from lists provided by UCSC Genome Browser's gene list.

Figure 3: Input file page for “Around CpG” utility.

For Around TSS Design.

Figure 4 shows the input file page for Around TSS utility. Here, the query file is Gene symbol list file. The gene symbols are taken from lists provided by UCSC Genome Browser's gene list. Other than uploading gene symbol list file and corresponding genome, it also requires special parameters i.e. Length of sequence upstream of TSS and Length of sequence downstream of TSS to pick query sequence around TSS. Both of these parameter values have been assigned with some default values for testing purpose.

Figure 4: Input file page for “Around TSS” Utility.

For Around max cut-site region Design.

Figure 5 shows the input file page for Around max cut-site region utility. Here, the query file is same as for regular primer design. Other than uploading query file and database, it also requires special parameters which are Number of Cut-sites, Cut-sites and Length of the Cut-site region to pick query sequence around region with maximum of those cut-sites. All these parameter values have been assigned with some default values for testing purpose.

Figure 5: Input file page for “Around max cut-site” utility

Step2: Input Parameters

Next stage of PRIMEGENS server is to provide all input parameters for primer design. All parameters have been set to some default values as standard parameters for best primer design. Input parameters on this page of the server are divided into five sections as follows:

1. *Algorithm Type*

In this, user can choose to design primers by three primer design algorithms supported by PRIMEGENS software or design just probes by choosing the last option. SSPD has been selected by default.

2. *Parameters required for Blast and Primer3 program*

Here, user can set parameters for MegaBLAST to look for cross hybridization of primers in database sequences provided by them. Then, for Primer3 parameters, user can provide specific desired characteristics of the primer that can be used by a third party program, Primer3 to design primers. For example, melting temperature, primer length, etc.

3. *Parameters required for Fragment Specific Primer Design (FSPD) program*

These parameters are used by PRIMEGENS only when it has to design primers using FSPD algorithm. Here, user can provide parameters for primer design only if they opted for algorithm type as FSPD.

4. *Parameters required for Probe Specific Primer Design (PSPD) program*

These parameters are used by PRIMEGENS only when it has to design primers using PSPD algorithm. Here, user can provide parameters for primer design only if they opted for algorithm type as PSPD.

5. *Parameters for Probe Design*

These parameters are used by PRIMEGENS only when it has to design only sequence specific probes. Here, user can provide parameters for probe design only if they opted for algorithm type as Probe Design.

Figure 6-12 below show the input parameter pages of PRIMEGENS tool. Here user can provide PRIMEGENS their own values or just run PRIMEGENS using all default values. Figure7 shows one of the help pop-ups available for each parameter by clicking the questionmark symbol in front of each.

PRIMEGENS-w3
High-throughput Primer and Probe Design

Digital Biology Laboratory
University of Missouri - Columbia

Home
Algorithm

- SSPD
- PSPD
- FSPD
- Check Binding Specificity
- Probe Design Only

Implementation
Database Support
References
Download User Manual

Input Parameters
(All parameters have been set to a default value.)

Design Algorithm

☒ SSPD: Sequence Specific Primer Design
 ☐ FSPD: Fragment Specific Primer Design
 Note: Useful for very long sequences.

☐ PSPD: Probe Specific Primer Design
 Note: FSPD does not work for Genomic Database.

☐ Only Probe Design

Next

Next page gives the user all parameter specifications used for primer/probe design.

Copyright | Acknowledgement | Contact Us

Designed by Garima Kushwaha,
Digital Biology Laboratory, MU Informatics Institute, University of Missouri - Columbia.

Figure 6: Page for setting algorithm type.

PRIMEGENS-w3
High-throughput Primer and Probe Design

Digital Biology Laboratory
University of Missouri - Columbia

Home
Algorithm

- SSPD
- PSPD
- FSPD
- Check Binding Specificity
- Probe Design Only

Implementation
Database Support
References
Download User Manual

Input Parameters
(All parameters have been set to a default value.)

Parameters required by Primer3 program to generate primers.

General Parameters

Primer min. Tm (default: 60.0)

Primer opt Tm (default: 60.0)

Primer max Tm (default: 66.0)

Primer opt size (default: 20)

Primer max size (default: 23)

Primer min GC (default: 40)

Primer GC clamp (default: 0)

Primer product opt Tm (default: 0.0)

Primer product min Tm (default: -1000000.0)

Primer product size range (default: 150-170)

Primer product opt size (default: 0)

Primer num Ns accepted (default: 0)

Primer liberal base (default: 0)

Primer first base index (default: 0)

Primer start codon position (default: -1000000)

Primer self end

Primer self any

PRIMER_MIN_TM

Minimum acceptable melting temperature(Celsius) for a primer oligo.

Figure 7: Page for setting Primer3 parameters for primer design.

After setting all Primer3 parameters and clicking “Next” button PRIMEGENS asks to set BLAST parameters. Figure 8 shows the page to set BLAST parameters.

PRIMEGENSw3
High-throughput Primer and Probe Design

Digital Biology Laboratory
University of Missouri - Columbia

Home
Algorithm

- SSPD
- PSPD
- FSPD
- Check Binding Specificity
- Probe Design Only

Implementation
Database Support
References
Download User Manual

Input Parameters
(All parameters have been set to a default value.)

Parameters required by BLAST program.

Word size [?](#) (default:11)
Gap Open Penalty [?](#) (default:4)
Gap Extension penalty [?](#) (default:2)

[Next](#) [Back](#)

[Copyright](#) | [Acknowledgement](#) | [Contact Us](#)

Designed by Garima Kushwaha,
Digital Biology Laboratory, MU Informatics Institute, University of Missouri - Columbia.

Figure 8: Page for setting BLAST parameters.

After setting all BLAST parameters and clicking “Next” button PRIMEGENS asks to set parameters specific to PRIMEGENS. Figure 9-12 shows the page to set these parameters.

PRIMEGENSw3
High-throughput Primer and Probe Design

Digital Biology Laboratory
University of Missouri - Columbia

Home
Algorithm

- SSPD
- PSPD
- FSPD
- Check Binding Specificity
- Probe Design Only

Implementation
Database Support
References
Download User Manual

Input Parameters
(All parameters have been set to a default value.)

Parameters required by PRIMEGENS algorithm.

Email for notification (optional)
print *Result Display [?](#)
Number of Primers [?](#) (default:10)
Allowed Hybridization in results [?](#) (default:0)
Maximum possible Amplicon Size [?](#) (default:9999)
Minimum Oligo size [?](#) (default:15)
print *Sequence in Bisulfite Mode [?](#) ☐
print *Number of CpG sites in primer [?](#) (default:2) Note: If uploaded sequences are in bisulfite mode.

[Back](#) [Run_PRIMEGENS!](#)

[Copyright](#) | [Acknowledgement](#) | [Contact Us](#)

Figure 9: Page for setting parameters specific to PRIMEGENS.

PRIMEGENS-w3
High-throughput Primer and Probe Design

Digital Biology Laboratory
University of Missouri - Columbia

Home
Algorithm

- SSPD
- PSPD
- FSPD
- Check Binding Specificity
- Probe Design Only

Implementation
Database Support
References
Download User Manual

Input Parameters
(All parameters have been set to a default value.)

Parameters for Fragment specific Primer Design (FSPD)

Fragment Length (default:500) Fragment Overlap (default:20)

Parameters required by PRIMEGENS algorithm.

Email for notification (optional)

print "Result Display" detailed

Number of Primers (default:10)

Allowed Hybridization in results (default:0)

Maximum possible Amplicon Size (default:9999)

Minimum Oligo size (default:15)

print "Sequence in Bisulfite Mode" ☐

print "Number of CpG sites in primer" (default:2) Note: If uploaded sequences are in bisulfite mode.

[Back](#) [Run_PRIMEGENS!](#)

[Copyright](#) | [Acknowledgement](#) | [Contact Us](#)

Figure 10: Page for setting parameters specific to PRIMEGENS when algorithm type as FSPD is chosen.

PRIMEGENS-w3
High-throughput Primer and Probe Design

Digital Biology Laboratory
University of Missouri - Columbia

Home
Algorithm

- SSPD
- PSPD
- FSPD
- Check Binding Specificity
- Probe Design Only

Implementation
Database Support
References
Download User Manual

Input Parameters
(All parameters have been set to a default value.)

Parameters for Probe specific Primer Design (PSPD)

Min. filter length (default:50) Max. similarity (default:0.75)

Min. probe length (default:100)

Parameters required by PRIMEGENS algorithm.

Email for notification (optional)

print "Result Display" detailed

Number of Primers (default:10)

Allowed Hybridization in results (default:0)

Maximum possible Amplicon Size (default:9999)

Minimum Oligo size (default:15)

print "Sequence in Bisulfite Mode" ☐

print "Number of CpG sites in primer" (default:2) Note: If uploaded sequences are in bisulfite mode.

[Back](#) [Run_PRIMEGENS!](#)

[Copyright](#) | [Acknowledgement](#) | [Contact Us](#)

Figure 11: Page for setting parameters specific to PRIMEGENS when algorithm type as PSPD is chosen.

High-throughput Primer and Probe Design University of Missouri - Columbia

Home
Algorithm

- SSPD
- PSPD
- FSPD
- Check Binding Specificity
- Probe Design Only

Implementation
Database Support
References
Download User Manual

Input Parameters
(All parameters have been set to a default value.)

Parameters for Probe Design (ONLY)

Probe BLAST hit (default:2)

Max. GC% (default:80)

Min. GC% (default:40)

Probe Length (default:100)

Probe Period (default:20)

Probe Region (default:200)

Unique ☐

Parameters required by PRIMEGENS algorithm.

Email for notification (optional)

print "Result Display" detailed

Number of Primers (default:10)

Allowed Hybridization in results (default:0)

Maximum possible Amplicon Size (default:9999)

Minimum Oligo size (default:15)

print "Sequence in Bisulfite Mode" ☐

print "Number of CpG sites in primer" (default:2) Note: If uploaded sequences are in bisulfite mode.

Back Run_PRIMEGENS!

Press RUN PRIMEGENS button for final execution of PRIMEGENS.

Figure 12: Page for setting parameters specific to PRIMEGENS when algorithm type as Probe Design is chosen.

After filling up all these parameter forms, user should hit “RUN PRIMEGENS” for the final run of the primer design program. User can hit “RUN PRIMEGENS”, without putting any value on this page and PRIMEGENS will design primers using all default parameters.

After running PRIMEGENS, server will show the link to find the output files. Figure13 shows the page with the link that comes after PRIMEGENS starts running.

PRIMEGENS-w3 Digital Biology Laboratory University of Missouri - Columbia

High-throughput Primer and Probe Design

Home
Algorithm

- SSPD
- PSPD
- FSPD
- Check Binding Specificity
- Probe Design Only

Implementation
Database Support
References
Download User Manual

Thank You for using PRIMEGENS.

All your input and outfiles will be available here <http://primegens.org/Garima/cgi-bin/8488> after the completion of primer design. PRIMEGENS is running. It might take few minutes to hours to update the final output.

Click "View Status" button to check PRIMEGENS run status. If PRIMEGENS run is complete, clicking this button will automatically redirect user to result page.

View Status

Return to our [home page](#).

Click on this link to access your final results. This link will show all the input and output files for the current process

Copyright | Acknowledgement | Contact Us

Designed by Garima Kushwaha,
Digital Biology Laboratory, MU Informatics Institute, University of Missouri - Columbia.

Figure13: Page after running PRIMEGENS.

Figure 14(a) shows the page that come on on hitting “View Status” button on page shown in Figure 13, if PRIMEGENS’ run is not finished. Figure14(b) shows confirmation pop-up that shows after pressing refresh button on its next page PRIMEGENS is still running for the job submitted. This absolutely safe to press “Resend” without losing design results and keep refreshing to check the PRIMEGENS’ completion. It takes few minutes for sample data for testing purpose.

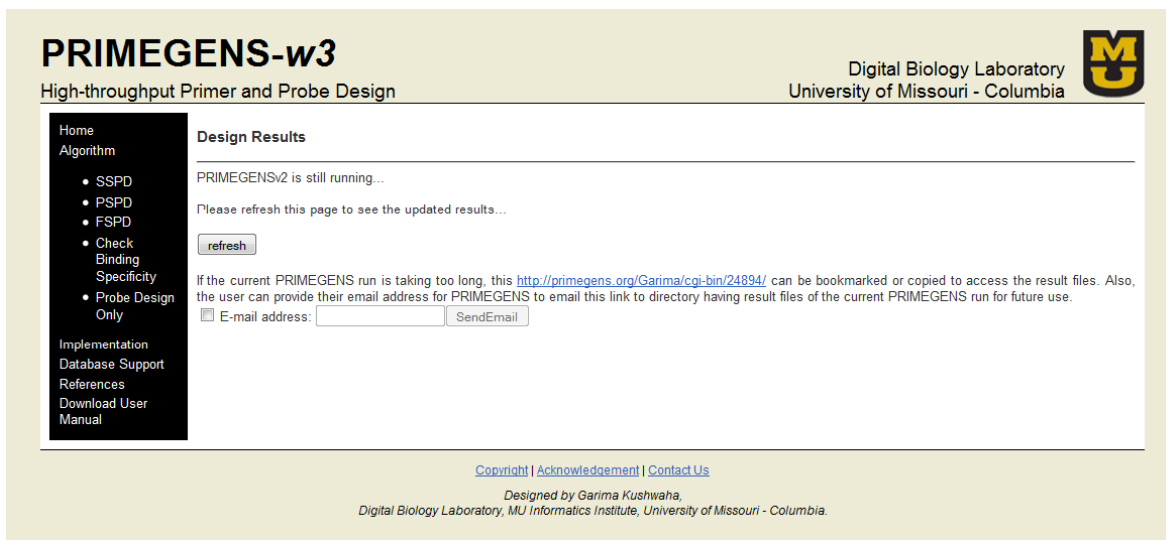


Figure 14(a): Next page after hitting “View Status” button on last page shown by Figure13.

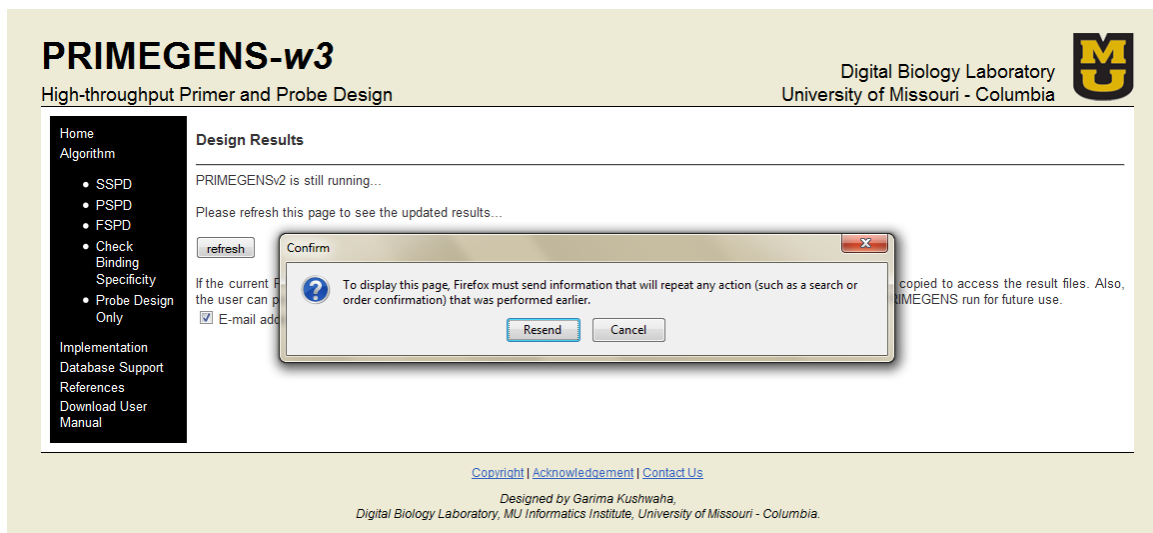


Figure 14(b): Confirmation pop-up on refreshing page.

Step3: Result Visualization.

On PRIMEGENS’ succedssfull execution and primer or probe design best results are shown in a form of table on web page with all information about each designed primers or probes, as shown in Figure 15(a). Double clicking on any row of this table or in other words each designed

primer record visualize the position of both left and right primer on its corresponding query sequence as shown in Figure 15(b). Also, name and links to all output files generated by PRIMEGENS are shown for user to see the results in their browser or right click and download them to their computer. All these files are still in the same directory as was provided in the link.

High-throughput Primer and Probe Design University of Missouri - Columbia

Home
Algorithm

- SSPD
- PSPD
- FSPD
- Check Binding Specificity
- Probe Design Only

Implementation
Database Support
References
Download User Manual

Design Results

Best primers

*Double click on any primer pair result in the following table to visualize their position on its corresponding sequence.

Query name	Left Primer-Sequence	Left primer-Start position	Left primer-Length	Left primer-Tm	Left primer-GC-content	Right primer-Sequence	Right primer-Start position	Right primer-Length
>S11529022 Zea ...	CGACATGGTCAAGGTCATCT	1145	20	58.5	50	AGTCATTTCGGGGTGAATCC	1313	20
>S22334604 Zea ...	AAACAGGGTAGAGCGTgagga	32	21	61.2	52	AAGTTGATGGGCACGATCTC	199	20
>S11417968 Zea ...	CACAAGTAGCGAAAtggcag	77	20	59.5	50	ACACTGTTGTCCTGTTCTCCT	244	22
>S11431057 Zea ...	GTTTGCAATTGTTGCGCTCC	14	20	63.9	50	TAACAAGTTGGCAACCCCT	164	20

Following are the result files available after running PRIMEGENS:

QueryFormat1_ZeaMays.fa	Query file
QueryFormat1_ZeaMays.fa_query_failed.txt	sequences with failed primer design
QueryFormat1_ZeaMays.fa_query_primer_list.txt	alternate primers for each sequence
QueryFormat1_ZeaMays.fa_primer.xls	best primer sequences

Figure 15(a): Primer Design result visualization.

Home
Algorithm

- SSPD
- PSPD
- FSPD
- Check Binding Specificity
- Probe Design Only

Implementation
Database Support
References
Download User Manual

Design Results

Best primers

*Double click on any primer pair result in the following table to visualize their position on its corresponding sequence.

Query name	Left Primer-Sequence	Left primer-Start position	Left primer-Length	Left primer-Tm	Left primer-GC-content	Right primer-Sequence	Right primer-Start position	Right primer-Length
>S11529022 Zea ...	CGACATGGTCAAGGTCATCT	1145	20	58.5	50	AGTCATTTCGGGGTGAATCC	1313	20
>S22334604 Zea ...	AAACAGGGTAGAGCGTgagga	32	21	61.2	52	AAGTTGATGGGCACGATCTC	199	20
>S11417968 Zea ...	CACAAGTAGCGAAAtggcag	77	20	59.5	50	ACACTGTTGTCCTGTTCTCCT	244	22
>S11431057 Zea ...	GTTTGCAATTGTTGCGCTCC	14	20	63.9	50	TAACAAGTTGGCAACCCCT	164	20

>S11417968 Zea mays heat shock protein 26 (HSP26) mRNA, complete cds; nuclear gene for plastid product /cds=p(91,813) /gb=L28712 /gi=453669 /ug=Zm.11 /len=946


TTTTTACCAAATCCGGACAGCTCTTGAGCTCTCTCAGTCTCAGCATTTTCAGTTCCAGATTACAGAGCCCCAGCCACACA
AGTAGCGAAAtggcagtgctccgttcgcgattgcccgcggctctcccCGTTGCGCGCCTCCCGTCCGCGCCTGGAGGCCGGCGC
ACGGGTTTGCCTGTCGCGGAGAGCCCGCTCGCTCGCCGTGCGATCCGCGCGCAGGAGAACAGGGACAACAGT
GGCGGCAACAGGCAGCAGGGCAACGCCGTGACGCGCCGCCCGCGCGCGACA
GCGCTGCATCTCCCATCCCGTTCGGATTAGTTGACCCGATGTCGCCATGCGGACGATGCGGCAGATGCTGGA
CACGATGGACAGGCTGTTGACGACGCCGTTGGGTTCCCATGGGACACGAGGTCCCGCGCACCAACGGCGA
CGTACGCTGCCGTGGGACATCGTGAGGACGAGAAGGAGTGGAAGATGCGGATCGACATGCCGGGCTCGCGCG
CGACGAGGTGAAGGTGATGGTGGAGGACGACACGCTGTCATCCGAGGGGAGCACAGAAGGAGGAGGGCGCGG
AGGGCGGACGCGGGGACGCGGACGCGGAGCGGTTGGTGAAGCAGCGCAGCGTGAGCTCTACGACATGCGGCTGGCG
CTGCCGATGAGTGCAGCAAGAGCAAGTGCAGGCGCGGAGCTCAAGAACGCGCTGCTCTGTCACCGTGCCTCAAGA
CCGAGGTGGAGCGCAAGGTCTACGAGTGCAGGTCCAGTAGACAACTGAGCTGTACGACGCGGTCTTTGTGTAGCA
CTAGCGTGGGTTGGGCTCTTCGCGCCGACCGTATCAGTACTCGTACCCACGAGTGGAAATATGTGAATAATAAAC
GCTTCAATTCTGCTATTCC

Figure 15(b): Visualizing Primer position in query sequence.

PRIMEGENS-w3

High-throughput Primer and Probe Design

Digital Biology Laboratory
University of Missouri - Columbia



Home
Algorithm

- SSPD
- PSPD
- FSPD
- Check Binding Specificity
- Probe Design Only

Implementation
Database Support
References
Download User Manual

Design Results

Query Name	Query Fragment (start-end)	Probe Sequence
S11529022	0000-0099	GCACGAGAACCGCTCCTCTGGTCCGTGGCCATGGCGTCGTCATGCTCTCCGCTACCACTGCCACTCCAGCAGGGGGGCGG
S11529022	0500-0599	CGGCATCGACCTCGTCATCGAGGGCAACGGCGCTCTTCGTCGACCGCGAGGGCGCGGGGAAGCACATCCAGGCGGGGGCCAAAG/
S11529022	1000-1099	AGGTGAACCAAGCGTTCCGCGACGCCCGGGCCAAACGAGCTCACGGGCATCCTCGAGGTCTGCGACGTGCCGCTCGTGTCCGTC
S22334604	0000-0099	GACGACGTTGCGCAGAGAATCCCCAAGCAAACAACAGGGTAGAGCGTgaggagaggaggaggaggaggaggaggTTGGGTCTGT
S22334604	0500-0599	TGAGAAGCTGAAGAAGGTGCTAGAGGTGTACGAGGCACGCCTGACCAAGTGCAAGTACCTTGCTGGAGACTTCTCAGCCTCGC
S11417968	0000-0099	TTTTTACCAAATCCGGACAGTCTTGAGCTCTCTCAGTCTCAGCATTCAGTCCAGATTACAGAGCCCCAGCCCCACAAGT

Following are the result files available after running PRIMEGENS:

[QueryFormat1_ZeaMays.fa](#) Query file
[QueryFormat1_ZeaMays.fa_probe.txt](#) Probe design results.

[Copyright](#) | [Acknowledgement](#) | [Contact Us](#)
 Designed by Garima Kushwaha,
 Digital Biology Laboratory, MU Informatics Institute, University of Missouri - Columbia.

Figure 16: Probe Design Result Visualization.

Input File Format

PRIMEGENSw3 supports query file in FASTA format. Following figure shows one sample file for this format.

```
>gi|2270989:16-516 Glycine max dehydrin (GmPM12) mRNA, complete cds
ATGGCTGAAGCACAACACGAGACCAGCATGGCAACCCTGTCCCACTACCGATCAATACGGTAATCCGG
TTATCTTAACGACGAGCGCGGTAATCCCGTCCAACCTCACTGGTGTGCTACCAACGCTACCGGCACAGC
AGGTTCTGGGTTTGGGTCTATGGTACCGGTGCTTACGGTGGTGGTGAAGTGAACCAACCGTTGCAGAT
CTTTGGCAACCAACCAAGGAGTGGCAGGGAGGCTAGAGAGCTTCGTCGTTCTCCAGTTCAAGCTCTA
GCTCGCTGAGGATGATGGGCAAGGTGGGAGGAGGAAGAAGGGAGTGAAGGATAAAATAAAGAGAAACT
ACCAGGGGTAGGAGGAGGAATAATAAAGGAGCATGCACACACAACACTGCTCAACCAACCGCCACT
AACCACCTGCTGATCAGCATGAGAAGAAGGGCATAATGGAGAGGATCAAGAAAAATTGCTGGCCACC
ACACCCACTGA
>gi|6648967:65-838 Glycine max seed maturation protein PM26 (PM26) mRNA, complete cds
ATGAGTCAGGAGCAGCCACGGCGTCCCAAAGGCCAAAACCCCATCAATACGGCGACGTTTTTGTGCTCT
CCGGCGACCTTGACAGAAAGCCCGTCGCACCGGAAGATGCTGCCATGATGCAAGCGCGGAACTCGAGT
GCTCGGACAAACCAACCCGGCGGAGCAGCTTCCGTCATGCAATCTGCCGCCACCAAGGAATGAACAGGCT
GGCCTTGTGCGTACCGGGACGTCACCGACGTTACCGGCGACCGTGGCGTCAAGTCAAGAACTAAAG
TCCCTGGAAGACGCATTATAACCGAGGCTGTTGGTGGCCAGGTTGTGGAGCAGTATGTGGAGGCAACTCC
GGTTGAGGCGAGGGCGAAGCAGTGCATTAAGGAGAATGCCATAACAATAGGAGAGGCATTGGAGGCGACG
GCACAGACTGTGGGTGAGAAAGCGGTGGATCAGAGTGACGCTTCGCGATTACGGCGCGGGAGGTGAGA
GCAACGGGGAAGTAAACGTTATAACGCCGGGTGGACTTGGCGCTATGGCTCAATCAGCTGCTGCTTATAA
TGCTGACTGCAAGCTTGACCAAGGCCAAGGTCAAGCTCGCCGACATTTTGGCCGGAGCCACAGCCAAGTTG
CCCGCGGACAAGGCCGCCACACTGCAAGATGCTGAAGGTGTAGCGTGTGCTGAGGTGAGGAACAACCTG
ATGCCACCGCCACTCCCGGTGGCGTAGCCGCTTCTGTTGCGGCTGCTGCTAGGCTCAATGAAATGTTAA
GTAA
>gi|4838146:81-503 Glycine max seed maturation protein PM30 (PM30) mRNA, complete cds
ATGGCATCCCATAGGCAAGCTATGAAGCTGGTCAAACTAAGGGCCGAAC TGAGGAAAAGACGAACCAGA
CGATGGGCAATATTGGAGAGAAGGCTCAAGCTGCAAGGAGAAGACCCAGGAAATGGCCCAAGCTGCAAA
GGAGAAGACCCAACAAACAGCCCAAGCTGCCAAGGACAAGACTTGCGACACTTCCCAAGCGGCAAGGAG
AAGACCAACAGAATACAGGAGCTGCTCAACAAAAGACCTCAGAGATGGGCGAGTCCACGAAGGAATCGG
```

For database file, PRIMEGENS support two types of format.

1. Single type database, i.e. one file containing all sequences in Fasta format (eg. Glycine max database).
2. Genome type database, whole genome in multiple files, i.e. one file per chromosome.

Currently, PRIMEGENS allow user to upload only single type database i.e. a single file with file size ~10MB. Web-server provides, in-house database for various model organisms, which user can select. In case user wants to use genome for any other organism they can contact PRIMEGENS developer with this request to for support.

In query file, user can input the nucleotide sequence for each query sequence or can just give gene names or chromosome position without their nucleotide sequence. In case nucleotide sequence is not provided and gene name is given, then database type should be single type database (mentioned in database drop down menu) or uploaded database sequence. But if chromosome position is given, database type should be genome type database where one file is present per chromosome.

Output format

Different number of and types of output files are generated by different design algorithm. All three primer design algorithms (SSPD- Sequence Specific Primer Design, FSPD – Fragment Specific Primer Design and PSPD – Probe Specific Primer Design) generate three different types of output files as follows:

1. Excel sheet: best primer pair

(Named as name of the query file followed by “primer.xls”)

This file contains best primer pair for each input query sequence along with other types of details, as follows:

Column Name	Description
QUERY_NAME	Name of the Query Sequence
LEFT_PRIMER	Left/Forward primer sequence
LEFT_PRIMER_START_POSITION	Start position of Left/Forward primer
LEFT_PRIMER_LENGTH	Length of Left/Forward primer
LEFT_PRIMER_TM	Melting temperature of Left/Forward primer
LEFT_PRIMER_GC_CONTENT	GC content of Left/Forward primer
RIGHT_PRIMER	Right/Reverse primer sequence
RIGHT_PRIMER_START_POSITION	Start position of Right/Reverse primer
RIGHT_PRIMER_LENGTH	Length of Right/Reverse
RIGHT_PRIMER_TM	Melting temperature of Right/Reverse
RIGHT_PRIMER_GC_CONTENT	GC content of Right/Reverse
PRODUCT_SIZE	Product or amplicon size
HYBRIDIZATION	Number of hybridization for the primer in database.

Figure- shows one sample of excel sheet output file generated by PRIMEGENS.

QUERY_NAME	LEFT_PRIMER	START	LEN	TM	GC	RIGHT_PRIMER	START	LEN	TM	GC	SIZE	HBDRN
>Glyma0070500210.1	TGGTGAGGAGGACTGAAAG	256	20	60.23	55	TGAAACCCAAAACTCCG	394	20	59.95	40	139	1
>Glyma01901750.1	AAGAGTGTGAAGCTTGGAT	831	20	60.02	50	CCATATCTCCAAATCCCT	924	20	59.97	50	94	1
>Glyma01904300.1	CAAGAGAACGGCCAAAGAG	1448	20	59.99	50	AAAGGGTGTGATCACTGG	1602	20	60.11	50	155	2
>Glyma01904300.2	CGAGTACAATCGCCAGACAA	820	20	59.86	50	TGCATGTCTTCTTGGAGTG	915	20	60.02	55	96	2
>Glyma01904750.1	AATGAAGGCAATGCCAATCTC	841	20	60.04	45	GGTCAGCTTGATGGAAAA	927	20	60.05	45	87	1
>Glyma01907600.1	CTTCGCCACTCTATCAAGC	158	20	59.98	55	GGAGTACCGAAGCTCGTTGT	308	20	60.04	55	151	1
>Glyma01909310.1	GTGAATGTCTTAAGGGGCAA	2621	20	59.93	50	ACAATGCCACAAGACCATGA	2725	20	59.97	45	105	1
>Glyma01927720.1	TGGGTTTATCCAGTTCCAG	332	20	59.78	50	CCTCTCTCTCAGATGGTC	476	20	59.95	60	145	1
>Glyma01930300.1	GTCTTCAAAAGGGATGGCAA	896	20	60.05	45	ATGACGGAGTTGGTGGAGAC	1026	20	59.97	55	131	1
>Glyma01936040.1	CCGCAGAGAAGGACAAAC	148	20	59.99	55	AGGGTAAAGCAACAGAGCGA	245	20	60.02	50	98	1
>Glyma01937090.1	CAATTTCCATATCCCAACGG	41	20	60.01	45	TATAGGCTGGATTGACGC	183	20	60.06	50	143	1
>Glyma01937760.1	ATCCCCCAGGAAAAAGAGA	4345	20	59.88	45	GCCTCTATGCCTATGGCTTC	4502	20	59.84	55	150	1
>Glyma01937760.2	AGGTGGGTGCTGTCAAAGTC	4569	20	60.16	55	AACAGCAGCAATGTTGAC	4667	20	59.92	45	99	1
>Glyma01939260.1	GCAACTCTCCGTTGAACCTC	684	20	59.85	55	AAGCGTTGTGTTTGTTC	791	20	60.02	45	168	1
>Glyma01939800.1	TGCAGAGAACATGGCTTCAG	138	20	60.14	50	AGGTCCGGGTGAGTCTCTTT	292	20	60.11	55	155	1
>Glyma01941850.1	GCCAACTGTGAGAACGCA	857	20	60.03	50	CACCTCTCCAGAGGACAC	969	20	59.99	60	113	2
>Glyma01941850.2	GCTGGCAATCAATACAGGT	1690	20	59.96	50	CCCAAACTGTCTCAACATT	1801	20	59.97	45	112	1
>Glyma01945410.1	ACATAGAGCGTGCAAACTGT	1564	20	59.94	50	CCATACAGGAATCGCAGGT	1644	20	59.96	50	81	1
>Glyma01945740.1	TCACACAGAGAATTACGCGG	64	20	59.86	50	CACCATTTCAAGCCAGTT	231	20	59.97	45	168	1
>Glyma01945740.2	GACCAGCTCAAGACAGC	240	20	60	55	CCAAAAGCATGGCAAGAT	337	20	60.07	40	98	2
>Glyma0202740.1	GCACTGATTTTACGCAGAA	133	20	60	45	ATCAGTGGCATCATGCTTCA	233	20	60.23	45	101	1
>Glyma0203400.1	AGCACGAGCTGGATTTGTTT	807	20	59.88	45	TGCACTGTCTTCTGGAGTG	924	20	60.02	55	118	2
>Glyma0203400.2	AGCACGAGCTGGATTTGTTT	807	20	59.88	45	TGCACTGTCTTCTGGAGTG	924	20	60.02	55	118	2
>Glyma0205670.1	TTCATAAAATCGGGTGGAGC	33	20	59.9	45	GTGTGAACGCGGATAGCAA	125	20	59.87	50	93	2

2. Alternate primer pairs (detailed)

(Named as name of the query file followed by "primers_list.txt")

This file contains alternate primer pairs for each input query sequences. In case user wants to select alternate primer pairs, this file provides multiple choices for selecting primer pairs for each query sequence. This file also contains the similar information as that in first file for every alternate primers.

```
>Glyma0070s00210.1
1) TGGTGAGGAGGAC TGAAG [ 256] TGAAGAC CAAAAC TCCG [ 394] psize 139 hrdn 1 Glyma0070s00210.1(129);
2) TGGTGAGGAGGAC TGAAG [ 256] ATCATCTGCAT TCTCGGGT [ 367] psize 112 hrdn 1 Glyma0070s00210.1(102);
3) ATGGTGAGGAGGAC TGAAG [ 255] TGAAGAC CAAAAC TCCG [ 394] psize 140 hrdn 1 Glyma0070s00210.1(130);
4) CCAGGGATGTGATTGATTC [ 600] TGACAGTTGGCAACAAATCC [ 747] psize 148 hrdn 1 Glyma0070s00210.1(138);
5) ATGGTGAGGAGGAC TGAAG [ 255] ATCATCTGCAT TCTCGGGT [ 367] psize 113 hrdn 1 Glyma0070s00210.1(103);
6) CGCAAGAGGGGTTGTGTAT [ 228] TGAAGAC CAAAAC TCCG [ 394] psize 167 hrdn 2 Glyma03g07770.1(157);Glyma0070s00210.1(157);
7) CGGAGTTTTGGGTTTTCA [ 375] CAAAAGGTCATCCGCAAT [ 470] psize 96 hrdn 2 Glyma03g07770.1(86);Glyma0070s00210.1(86);
8) AAAGAGACGCTGAAGCCAA [ 167] ATACACACCCCTTTTGGG [ 247] psize 81 hrdn 2 Glyma03g07770.1(71);Glyma0070s00210.1(71);
9) CAAGAAAGCCTATTCGAAG [ 109] GAATTTGGCTTCAGCGCTC [ 190] psize 82 hrdn 2 Glyma03g07770.1(72);Glyma0070s00210.1(72);
10) CGCAAGAGGGGTTGTGTAT [ 228] ATCATCTGCAT TCTCGGGT [ 367] psize 140 hrdn 2 Glyma03g07770.1(130);Glyma0070s00210.1(130);

>Glyma01g01750.1
1) AAGAGTGTGAAGCTGCGAT [ 831] CCATATCTCCAAATCCCT [ 924] psize 94 hrdn 1 Glyma01g01750.1(84);
2) AAGAGTGTGAAGCTGCGAT [ 832] CCATATCTCTCAATCCCT [ 924] psize 93 hrdn 1 Glyma01g01750.1(83);
3) AAGAGTGTGAAGCTGCGAT [ 831] AATCCAGCAC TGGCATA TCC [ 936] psize 106 hrdn 1 Glyma01g01750.1(96);
4) AAGAGTGTGAAGCTGCGAT [ 832] AATCCAGCAC TGGCATA TCC [ 936] psize 105 hrdn 1 Glyma01g01750.1(95);
5) CGGCAGGATTGAGAAATAA [ 774] AATCCAGCAC TGGCATA TCC [ 936] psize 163 hrdn 1 Glyma01g01750.1(153);
6) CCTGCGATTGAGGAGAGAG [ 843] CCATATCTCCAAATCCCT [ 924] psize 82 hrdn 1 Glyma01g01750.1(72);
7) CAGTGTGATTCGGATTTT [ 925] CCTCACTCAAAGGGATTCA [ 1082] psize 158 hrdn 1 Glyma01g01750.1(148);
8) GGATATGGCAGTGTGATTT [ 917] CCTCACTCAAAGGGATTCA [ 1082] psize 166 hrdn 1 Glyma01g01750.1(156);
9) TATTTGATGTGATGAGGCA [ 1304] CAAGATGCGCCATCTCAGA [ 1395] psize 92 hrdn 1 Glyma01g01750.1(82);
10) AGGGGATTGTGAGGATATG [ 905] TAAATCTCGGCATTCATC [ 1036] psize 132 hrdn 1 Glyma01g01750.1(122);

>Glyma01g04300.1
1) CAAGAGACCGCCAAAGAG [ 1448] AAAGGGTGTGATCAACTGG [ 1682] psize 155 hrdn 2 Glyma01g04300.2(190);Glyma01g04300.1(145);
2) CGAGTACAAATCCGACAGAA [ 820] TGCACGTCTTCTGGAGTG [ 915] psize 96 hrdn 2 Glyma01g04300.2(86);Glyma01g04300.1(86);
3) TTAAGAGGAAGCCTTGCCA [ 1626] AAAGGGGGGAAGGGATTAT [ 1753] psize 128 hrdn 2 Glyma01g04300.2(118);Glyma01g04300.1(118);
4) TTAAGAGGAAGCCTTGCCA [ 1626] TCATTTTGGCATGCTTGAG [ 1713] psize 88 hrdn 2 Glyma01g04300.2(78);Glyma01g04300.1(78);
5) ACAAGAGACGGCCAAAGAA [ 1447] AAAGGGTGTGATCAACTGG [ 1682] psize 156 hrdn 2 Glyma01g04300.2(191);Glyma01g04300.1(146);
6) CCAGTTGATCAGACCTTTT [ 1583] TTTTGGCATGCTTGAGTGAC [ 1709] psize 127 hrdn 2 Glyma01g04300.2(117);Glyma01g04300.1(117);
7) CCAGTTGATCAGACCTTTT [ 1583] TCATTTTGGCATGCTTGAG [ 1713] psize 131 hrdn 2 Glyma01g04300.2(121);Glyma01g04300.1(121);
8) GGCCTTGAGGCTGTGAAATC [ 544] GCTCTTCCAAACAGTTGCG [ 689] psize 146 hrdn 2 Glyma01g04300.2(136);Glyma01g04300.1(136);
9) GACCATTCGACCATCTTCAT [ 705] ACTTGCTTTTGTCTGGCGAT [ 847] psize 143 hrdn 2 Glyma01g04300.2(133);Glyma01g04300.1(133);
10) GACCATTCGACCATCTTCAT [ 705] CCAGCTTGCTGCTCTCTTC [ 809] psize 105 hrdn 2 Glyma01g04300.2(95);Glyma01g04300.1(95);
```

3. Failed sequences

(Named as name of the query file followed by "query_failed.txt")

This file contains input query sequences in fasta format, for which primer design is failed. That is no primer pair found in the given constraints. User can use this file for primer design using PRIMEGENS with different primer design parameters.

In addition to these three files, PSPD generate an additional output file

1. Gene-specific fragment (only PSPD)

(Named as name of the query file followed by "query_failed.txt")

This file is generated only during Probe-specific primer design (PSPD). This file contains gene-specific fragment (probe) for each input query sequence that PSPD find using global alignment of query sequence with the database sequences. These are the gene-specific fragments that PSPD ultimately use to design primers for their corresponding query sequence. This file could be useful for microarray probe design. The primer pair designed for each query sequence as designed to amplify these gene-specific probes. This is a normal FASTA formatted file.